

Parallel Linear Search with no Coordination for a Randomly Placed Treasure*

Amos Korman¹ and Yoav Rodeh²

¹CNRS and University Paris Diderot., Paris, France, amos.korman@irif.fr

²Weizmann Institute of Science, Rehovot, Israel, yoav.rodeh@gmail.com

April 19, 2016

Abstract

In STOC'16, Fraigniaud et al. consider the problem of finding a treasure hidden in one of many boxes that are ordered by importance. That is, if a treasure is in a more important box, then one would like to find it faster. Assuming there are many searchers, the authors suggest that using an algorithm that requires no coordination between searchers can be highly beneficial. Indeed, besides saving the need for a communication and coordination mechanism, such algorithms enjoy inherent robustness. The authors proceed to solve this linear search problem in the case of countably many boxes and an adversary placed treasure, and prove that the best speed-up possible by k non-coordinating searchers is precisely $\frac{k}{4}(1+1/k)^2$. In particular, this means that asymptotically, the speed-up is four times worse compared to the case of full coordination.

We suggest a natural variant of the problem, where the treasure is placed uniformly at random in one of a finite, large, number of boxes. We devise non-coordinating algorithms that achieve a speed-up of $6/5$ for two searchers, a speed-up of $3/2$ for three searchers, and in general, a speed-up of $k(k+1)/(3k-1)$ for any $k \geq 1$ searchers. Thus, as k grows to infinity, the speed-up approaches three times worse compared to the case of full coordination. Moreover, these bounds are tight in a strong sense as no non-coordinating search algorithm for k searchers can achieve better speed-ups. We also devise non-coordinating algorithms that use only logarithmic memory in the size of the search domain, and yet, asymptotically, achieve the optimal speed-up. Finally, we note that all our algorithms are extremely simple and hence applicable.

*This work has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No 648032).

1 Introduction

Parallel algorithms are often evaluated by their *speed-up*, which measures how much faster the parallel algorithm with k processors runs in comparison to the best running time that a single processor can achieve. In [13], the authors focused on parallel *non-coordinating* algorithms. In such an algorithm, all processors operate independently, by executing the same protocol, differing only in the outcome of the flips of their random coins. A canonical example is the case of multiple random walkers that search in a graph [1]. Although most problems cannot be efficiently parallelized without coordination, when such parallelization can be achieved, the benefit can potentially be high. Particularly, as argued in [13], the gain can be not only in terms of saving in communication and overhead in computation, but also in terms of robustness.

One class of fundamental problems that enjoys a large non-coordination ratio in some circumstances, is the class of search problems over totally ordered sets. The objective of such *linear search* problems is to find a solution among a set of candidate solutions that are linearly ordered according to their quality. For instance, searching for a proper divisor of a random integer n is an illustration of linear search. In this case, enumerating the candidate divisors in increasing order, and checking them one after another, is the typical strategy to solve the problem, as the probability that n is divisible by a given prime is inversely proportional to this prime. Similarly, in cryptography, an exhaustive search attack is better proceeded by systematically checking smaller keys than longer ones, since checking a key typically requires exponential time in the key size. In general, linear search appears in contexts in which the search space can be ordered in a way such that, given that the previous trials failed, the next candidate according to the order is either the most likely to be valid, or most preferable, or the easiest to check.

One basic linear search problem is the *treasure-hunt* problem. There is one treasure hidden in one box B_x out of a linearly ordered set of boxes $\{B_i\}$. Searchers are unaware of the index x and their goal is to find the treasure as fast as possible. At each time step, a protocol executed by a searcher has the ability to peek into one box and see whether the treasure is present or not in that box. The protocol is terminated once one of the searchers finds the treasure. Boxes are listed in an order that reflects the importance of finding the treasure in a given box. That is, finding the treasure hidden in B_i for small values of i is more urgent than for large values of i . In the case of a solo searcher, the best strategy is to open the boxes in increasing order. In this case, the searcher will find the treasure in x time. Hence, for a given algorithm A , we measure the *speed-up* function of k searchers with respect to x as:

$$speedup_A(k, x) = \frac{x}{\mathbb{E}[\text{time to find } x \text{ with } k \text{ searchers running } A]}$$

where the expectation is taken with respect to the randomness of the searchers. Note that if coordination is allowed, one could simply divide the set of boxes evenly between the searchers leading to a speed-up of k . However, as mentioned in [13], as simple as this algorithm is, it is very sensitive to faults of all sorts. For example, if one searcher crashes at some point in time during the execution then the searchers may completely miss the treasure, unless the protocol employs some mechanism for detecting such faults. This observation motivates the study of non-coordinating algorithms.

In [13] the authors considered the case in which an adversary places the treasure in one of infinitely countable boxes. In this case, for any non-coordinating algorithm, the adversary can always find a box for which the speed-up would be at most 1. Therefore, the speed-up of an algorithm was defined as $\limsup_{x \rightarrow \infty} speedup_A(k, x)$. They then present Algorithm 1 that is very simple and achieves optimal performance, a speed-up of $\frac{k}{4}(1 + 1/k)^2$ for each k . Specifically, the speed-up is $9/8$ for two searchers, $4/3$ for three searchers, and roughly $k/4$ for k searchers as k grows larger. It is shown that no non-coordinating search algorithm can achieve better speed-ups.

In this paper, we expand on the techniques and results of [13], and solve a natural variant of the treasure-hunt problem. In our scenario, the treasure is placed *at random* in one of M boxes. We therefore define the speed-up of an algorithm as:

$$speedup_A(k) = \mathbb{E}[speedup_A(k, x)]$$

Algorithm 1 The optimal algorithm in [13] for the adversary placed treasure.

```

for  $i = 1$  to  $\infty$  do
    check a uniformly chosen box of those unchecked in  $\{1, \dots, i * (k + 1)\}$ 
    check another uniformly chosen box of those unchecked in the same range

```

where the expectation is taken with respect to the placement of x . The speed-up of k agents on M boxes is defined as the speed-up of the best algorithm for M boxes. That is:

$$\text{speedup}_M(k) = \sup\{\text{speedup}_A(k) \mid A \text{ works on } M \text{ boxes}\}.$$

Finally, the general speed-up of k agents is defined as the speed-up of the best algorithm for M boxes, where M is taken to infinity. That is:

$$\text{speedup}(k) = \liminf_{M \rightarrow \infty} \text{speedup}_M(k).$$

1.1 Our Results

We suggest a general framework to tackle the randomly placed treasure hunt problem and solve it completely for the case of a treasure uniformly placed in one of finitely many boxes.

Specifically, we prove that the optimal speed-up for non-coordinating algorithms running with $k \geq 1$ searchers is:

$$\text{speedup}(k) = \frac{k(k+1)}{3k-1}$$

This means that the best speed-up that can be achieved is $6/5$ for two searchers, $3/2$ for three searchers, and roughly $k/3$ for k searchers, as k grows larger. We stress that these bounds are tight in a strong sense as no non-coordinating search algorithm for k searchers can achieve better speed-ups, for any k .

Our results indicate that the best possible speed-up for the random setting is strictly higher than the best possible speed-up for the adversarial setting. Specifically, as the number of searchers grows, the multiplicative gap between the settings approaches $4/3$.

The non-coordinating algorithm achieving¹ the aforementioned ratio for the random setting with k searchers is so simple that it can be described in just a few lines (Algorithm 2). Note that for each k and M the algorithm is different. It is interesting to note that this algorithm finds a way to balance between

Algorithm 2 The asymptotically optimal algorithm.

```

for  $i = 1$  to  $M/k$  do
    check a uniformly chosen box from those unchecked in  $\{1, \dots, i * k\}$ 
while there are unchecked boxes do
    check a uniformly chosen box from those unchecked

```

two opposing forces: it gives higher probability of quickly opening the more important boxes, and yet tries to minimize as much as possible the overlap between different agents.

Despite the simplicity of Algorithm 2, finding it was not an easy task. Indeed, there are many possible algorithms that achieve an $O(k)$ speed-up, but coming up with an algorithm that has an exact optimal bound is not trivial. This is especially challenging as every slight change in the algorithm may significantly complicate its analysis.

Algorithm 2 achieves optimal speed-up for each k but requires $\Omega(M)$ bits of memory to store the elements that have already been checked. To circumvent this, we suggest and analyse a natural variant of this

¹Note that this algorithm achieves the bound in an asymptotic manner, meaning that as the number of boxes goes to infinity, the speed-up of the algorithm approaches the desired bound. For small number of boxes there might be better algorithms.

algorithm, which works in exactly the same fashion, except that it does not remember the boxes it already checked. It therefore requires only $O(\log M)$ bits of memory, and yet its speed-up is $k/3$, which is practically the same as that of the optimal algorithm for large k .

Concerning the technical aspects, in a similar fashion to [13], we first observe that the crucial aspects of non-coordinating algorithms can be represented by matrices and then approximate matrices by continuous functions. To show the upper bound on the speed-up we had to upper bound the integral of functions satisfying certain restrictions. As in [13], the biggest technical challenges lie in computing such integrals. This task turned out to be more difficult than in the adversarial case of [13], and we were therefore unable to directly compute these integrals. Instead, to overcome this, we needed two new tools: We extend a lemma of [13] to work on unbounded functions, and introduce the intuitive notion of the “zooming” of a function. Using this two new results, we can reduce the problem to solving simpler integrals. We then turn to analyze Algorithm 2 to lower bound the speed-up for k agents. One of the problems arising in the analysis of Algorithm 2 which did not occur in the limsup adversarial setting of [13], is that the first boxes cannot be ignored, as they contribute significantly to the average case analysis. To overcome this we prove a small, yet useful, result regarding the Gamma function.

As mentioned, we fully characterize the case of a treasure placed uniformly at random. We note, however, that the general framework of analysis presented in this paper could be useful in order to derive speed-up bounds with respect to other distributions as well. Handling other distributions, however, would require solving the corresponding particular integrals.

To sum up, our upper bound on the speed-up of non-coordination algorithms implies that there is a price to be paid for the absence of coordination, which is asymptotically a factor of three away from an ideal optimal algorithm that performs with perfect coordination. On the other hand, this price is actually reasonably low, and so in faulty contexts in which coordination between the searchers may yield severe overheads, it might well be worth giving up on coordination, and simply run our non-coordinating algorithm.

1.2 Related Work

Most of the parallel search literature deals with mobile agents searching graphs of different topologies. One example is search by multiple random walkers. In a series of papers [1, 5, 8, 7] several results regarding hitting time, cover time and mixing times are established, such as a linear speed-up for several graph families including expanders and random graphs.

Another classical example of mobile search is the linear search problem [3, 4] (and later [14]), where a treasure is hidden on the real line according to a known probability distribution. A searcher starts from the origin and wishes to discover the treasure in minimal expected time. This problem was reintroduced by computer scientists as the *cow-path* problem [2, 15]. There, similarly to this paper, the time complexity is measured as a function of the distance of the treasure from the starting point.

Motivated by applications to central search foraging by desert ants, the authors in [11, 12] considered the ANTS problem, a variant of the cow-path problem on the grid, and showed that a speed-up of $O(k)$ can be achieved with k independent searchers. Several variants of the cow-path problem and the ANTS problem were studied in [6, 9, 10, 16, 17, 18].

The major difference between our setting and the mobile agent setting, is that we allow “random access” to the different boxes. That is, our searcher can jump between different boxes at no cost, unlike the case of mobile agents on a graph which can only move from a vertex to a neighboring vertex in one time step.

2 Preliminaries

In this section we go over the setup of [13] with some minor changes that reflect our changed definition of speed-up. We then proceed to introduce some new tools to help us deal with the case of randomly placed treasure.

2.1 Setup

Our universe contains boxes indexed by $i \in \{1 \dots M\}$, and a treasure that is placed uniformly at random in one of them. At each time step, a searcher can peek into exactly one box. There are k searchers and they are all completely identical in that they have the same algorithm, yet their randomness is independent. They do not communicate at all, and their mutual purpose is to maximize the speed-up by minimizing the expected time until the first one of them finds the treasure. In our technical discussion it will be often easier to work with the *inverse* of the speed-up. Specifically, for an algorithm A and an index x , define:

$$\theta_A(k, x) = \frac{1}{x} \mathbb{E}[\text{time to find } x \text{ with } k \text{ searchers running } A]$$

where the expectation is taken over the randomness of the searchers. Taking expectation over the placement of x , we define:

$$\theta_A(k) = \mathbb{E}[\theta_A(k, x)]$$

So, an algorithm with $\theta(3) = 1/2$ means that running this algorithm on three searchers will result in an expected running time that is twice as fast as the trivial one-searcher algorithm.

2.2 From Algorithms to Matrices

Given an algorithm, we define the matrix N , where $N(x, t)$ is the probability that the algorithm executed with just one searcher has *not* visited box x up to (and including) step t . For example, the following matrix corresponds to Algorithm 2, where $k = 2$ and $M = 6$ (note that we stop after time $t = 6$ since afterwards, the matrix is all 0's).

$\begin{smallmatrix} t \rightarrow \\ x \downarrow \end{smallmatrix}$	0	1	2	3	4	5	6	...
1	1	1/2	1/3	1/4	1/6	1/12	0	
2	1	1/2	1/3	1/4	1/6	1/12	0	...
3	1	1	2/3	1/2	1/3	1/6	0	
4	1	1	2/3	1/2	1/3	1/6	0	...
5	1	1	1	3/4	1/2	1/4	0	
6	1	1	1	3/4	1/2	1/4	0	...

Some observations:

1. The value of $N(x, t)$ is $N(x, t - 1)$ multiplied by the probability that x will not be checked at time t . This is how we arrived at the matrix above.
2. The sum of row x is the expected time until the algorithm peeks into box x . Indeed, let $I_{x,t}$ denote the indicator random variable that is 1 iff $t < \text{the visit time of } x$. The sum of these over t is the visit time. Also, $\Pr[I_{x,t} = 1] = N(x, t)$, so we get the result by linearity of expectation. This means:

$$\theta_N(1, x) = \frac{1}{x} \sum_{t=0}^{\infty} N(x, t)$$

3. Given the matrix N for one searcher, what would be the N matrix for k searchers? The probability of x not being looked into up to step t is the probability that all k searchers didn't peek into it, which is $N(x, t)^k$. So by the same reasoning as the last point, we get:

$$\theta_N(k, x) = \frac{1}{x} \sum_{t=0}^{\infty} N(x, t)^k \tag{1}$$

4. Since $1 - N(x, t)$ is the probability that box x was peeked into by step t , summing these numbers over column t , we get the expected number of boxes checked by this time, which is of course at most t . Denoting:

$$C_N(t) = \sum_{x=1}^M 1 - N(x, t)$$

The *column requirement* is that for all t , $C_N(t) \leq t$. The matrix resulting from an algorithm will always satisfy the column requirement.

2.3 One Searcher

A simple and intuitive result is when considering just one agent. We show it here for completeness and also as intuition, as it is a very simple version of the proof for the general case.

Theorem 1. $\text{speedup}(1) = 1$.

Proof. The algorithm A that opens one box after another will take time x to find the treasure if it is in box x , and so: $\text{speedup}_A(1, x) = x/x = 1$, meaning that $\text{speedup}_A(1) = 1$. So, $\text{speedup}(1) \geq 1$.

On the other hand, take some algorithm and denote its matrix by N .

$$\theta_N(1) = \frac{1}{M} \sum_{x=1}^M \frac{1}{x} \sum_{t=0}^{\infty} N(x, t) = \frac{1}{M} \sum_{t=0}^{\infty} \sum_{x=1}^M \frac{1}{x} N(x, t) \quad (2)$$

We aim to show that $\theta_N(1) \geq 1$. For this purpose, we actually show that this inequality holds for every N taking values in $[0, 1]$ satisfying the column requirement. That is, we know that for every t ,

$$C_N(t) = \sum_{x=1}^M 1 - N(x, t) \leq t$$

Wanting to minimize (2) under this restriction means that we can deal with each column separately. So the question is, given some integer t , what are the values of $f_1, \dots, f_M \in [0, 1]$ that minimize $\sum_{x=1}^M f_x/x$ under the restriction that $\sum_{x=1}^M (1 - f_x) \leq t$. The answer is $f_x = 0$ for all $x \leq t$ and 1 for the others. Otherwise there is some $y > x$, s.t. one can decrease f_x a little and increase f_y by the same amount. This will result in a smaller sum without violating the constraint.

The resulting optimal matrix is therefore $N(x, t) = 0$ for all $x \leq t$ and 1 otherwise. This is exactly the matrix of the trivial algorithm, and so the optimal speedup is at most 1. \square

2.4 From Matrices to Functions

It will be much easier to work with a continuous version of our problem. For an interval $X \subseteq \mathbb{R}^+$ denote: $\mathcal{F}(X) = \{N : X \times [0, \infty] \rightarrow [0, 1] \mid N \text{ is continuous}\}$. For an $N \in \mathcal{F}(X)$, we say that N satisfies the column requirements if for all t :

$$C_N(t) = \int_X 1 - N(x, t) dx \leq t$$

We call such a function valid, and denote by $\mathcal{V}(X)$ the set of all such valid functions. We define:

$$\theta_N(k) = \frac{1}{|X|} \int_X \int_0^\infty \frac{1}{x} N(x, t)^k dt dx = \frac{1}{|X|} \int_0^\infty \int_X \frac{1}{x} N(x, t)^k dx dt$$

where we can change the integral because of Tonelli's theorem. The following claim shows a connection between matrices and functions:

Claim 1. For any algorithm A that works on M boxes, there is a function $N \in \mathcal{V}([1, M+1])$ such that for all k , $\theta_N(k) \leq \theta_A(k)$.

The proof is in Appendix A.1, where we choose N to be a continuous version of the step function derived from A 's matrix.

Claim 1 shows that upper bounding the speed-up of all functions in $\mathcal{V}([1, M+1])$ will upper bound the speed-up of all algorithms on M boxes.

3 Zooming in (or out) a Function

We introduce a new and important tool that will be used a couple of times in what proceeds. If we have some algorithm that works for M boxes, we would like to be able to use it on $2M$ boxes. Slowing it down by a factor of two, and using the extra probability to cover the extra boxes can do the trick. How to do this for algorithms is not clear, but for continuous functions it is simple and elegant.

Definition 1. Given some $N \in \mathcal{F}(X)$ and some $a, b > 0$, we define the zooming of N by (a, b) as:

$$N_{\vec{a,b}}(x, t) = N(x/a, t/b)$$

Where $N_{\vec{a,b}}(x, t) \in \mathcal{F}(aX)$.

The intuitive meaning of it is that we expanded the algorithm to work on a domain of size a times the original one, and slowed it down by a factor of b . What happens to the column integrals and to the speed-up?

Claim 2. For $N \in \mathcal{F}(X)$ and $a, b > 0$, we have $\theta_{N_{\vec{a,b}}}(k) = \frac{b}{a}\theta_N(k)$ and for all values t , $C_{N_{\vec{a,b}}}(t) = aC_N\left(\frac{t}{b}\right)$.

Proof. The inverse speed-up:

$$\begin{aligned} \theta_{N_{\vec{a,b}}}(k) &= \frac{1}{|aX|} \int_{aX} \left(\int_0^\infty \frac{1}{x} N_{\vec{a,b}}(x, t)^k dt \right) dx = \frac{1}{a|X|} \int_{aX} \left(\int_0^\infty \frac{1}{x} N(x/a, t/b)^k dt \right) dx \\ &= \frac{1}{a|X|} \int_{aX} \left(b \int_0^\infty \frac{1}{x} N(x/a, t)^k dt \right) dx = \frac{1}{a|X|} \int_X \left(ba \int_0^\infty \frac{1}{ax} N(x, t)^k dt \right) dx \\ &= \frac{b}{a} \cdot \frac{1}{|X|} \int_X \left(\int_0^\infty \frac{1}{x} N(x, t)^k dt \right) dx = \frac{b}{a} \theta_N(k) \end{aligned}$$

And the column integrals:

$$C_{N_{\vec{a,b}}}(t) = \int_{aX} 1 - N_{\vec{a,b}}(x, t) dx = \int_{aX} 1 - N(x/a, t/b) dx = a \int_X 1 - N(x, t/b) dx = aC_N\left(\frac{t}{b}\right)$$

□

Corollary 1. For $N \in \mathcal{V}(X)$ and any $a > 0$, we have that $N_{\vec{a,a}} \in \mathcal{V}(aX)$ and has exactly the same speed-up as N .

Proof. By Claim 2 and the fact that N satisfies the column requirements, $C_{N_{\vec{a,a}}}(t) = aC_N(t/a) \leq a(t/a) = t$. By the same claim we see that the speedup remains unchanged. □

4 Upper Bound

Theorem 2. *For all $k \geq 1$:*

$$\text{speedup}(k) \leq \frac{k(k+1)}{3k-1}$$

To prove the theorem, we show that for every $\delta > 0$, for large enough M , any algorithm A working on M boxes satisfies:

$$\text{speedup}_A(k) \leq (1 + \delta) \frac{k(k+1)}{3k-1}$$

The case $k = 1$ was dealt with in Section 2.3. Claim 1 says that if we prove this for all $N \in \mathcal{V}([1, M+1])$ then we have proved it on all algorithms designed for M boxes. Corollary 1 then says that by zooming N by $M+1$, proving the following claim will prove Theorem 2:

Claim 3. *There is some real function f , s.t. $\lim_{\epsilon \rightarrow 0} f(\epsilon) = 1$, and every $N \in \mathcal{V}([\epsilon, 1])$ satisfies:*

$$\theta_N(k) \geq f(\epsilon) \cdot \frac{3k-1}{k(k+1)}$$

We use a lemma² of [13]:

Lemma 1. *Fix $\epsilon \geq 0$ and $T \geq 0$. For continuous functions $a : [\epsilon, 1] \rightarrow \mathbb{R}^+$ and $f : [\epsilon, 1] \rightarrow [0, 1]$ where $\int_{\epsilon}^1 1 - f(x)dx \leq T$, the minimum of $\int_{\epsilon}^1 a(x)f(x)^k dx$, ranging over all possible f 's is achieved when*

$$f(x) = \min \left(1, \alpha a(x)^{-\frac{1}{k-1}} \right)$$

where α is a function of $a(\cdot)$ and T , and independent of x .

This lemma is exactly what we need to prove Claim 3. Informally, considering $N \in \mathcal{V}([\epsilon, 1])$, the column requirement is the condition on f in the lemma, just setting $T = t$, and by plugging in $a(x) = 1/x$, we get the form of the optimal function for each of the columns. For each t , we then find the maximal α solving $\int_{\epsilon}^1 1 - \min(1, \alpha x^{1/(k-1)})dx \leq t$, and thus get the optimal function N for a specific ϵ . We can then calculate $\theta_N(k)$ and take the limit as ϵ goes to 0 (that is, M goes to infinity) to get our result.

The only problem with this approach is that the integrals become difficult to solve when we start them at some arbitrary ϵ . We therefore take $\epsilon = 0$. Then, as we shall see, the integrals are pretty straightforward. Our plan is therefore as follows:

1. Extend Lemma 1. This is because it doesn't work now, as $a(x) = 1/x$ is not defined on 0, and cannot be extended in a continuous way to it.
2. Use the extended lemma to find the exact optimal function in the case of $\epsilon = 0$, and calculate its speed-up.
3. Use the result to come back to $\epsilon > 0$ with asymptotically insignificant change in speed-up.

This is the extended lemma:

²It is slightly changed here: the upper bound of the domain is 1 and not infinity. This changes nothing in the original proof and actually simplifies the statement of the lemma, as boundedness and integrability of $a(x)$ are immediate. We also change the requirement to inequality, and this again, changes nothing, as when the optimal solution is not 0, it will clearly take the equality.

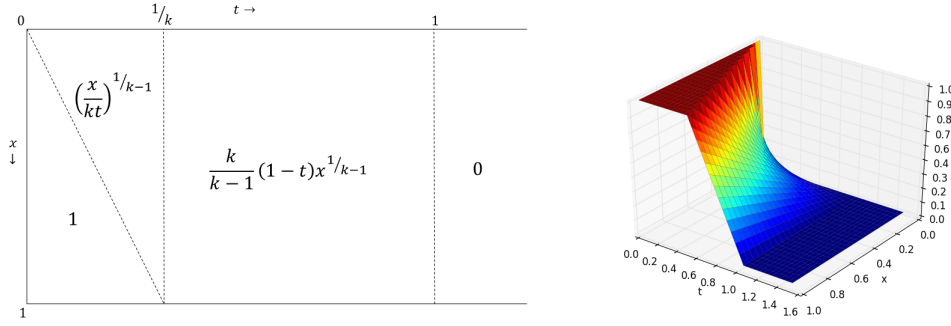


Figure 1: The optimal function OPT_k .

Lemma 2. Fix $T \geq 0$. For continuous functions $a : (0, 1] \rightarrow \mathbb{R}^+$ and $f : (0, 1] \rightarrow [0, 1]$ where $\int_0^1 a(x)^{-1/(k-1)} dx$ is defined and $\int_0^1 1 - f(x) dx \leq T$, the minimum of $\int_0^1 a(x)f(x)^k dx$, ranging over all possible f 's is achieved when:

$$f(x) = \min \left(1, \alpha a(x)^{-\frac{1}{k-1}} \right)$$

Where α is a function of $a(\cdot)$ and T , and independent of x .

It is proved in Appendix A.2 by using the original lemma on subsets $[s, 1]$, taking the limit of the resulting functions, and then using the dominated convergence theorem.

As we can see the condition on a is a bit bizarre, but it proves what we need:

Corollary 2. Among all continuous functions $f : (0, 1] \rightarrow [0, 1]$ satisfying $\int_0^1 1 - f(x) dx \leq T$, the one that minimizes $\int_0^1 \frac{1}{x} f(x)^k dx$ is $f(x) = \min \left(1, \alpha x^{\frac{1}{k-1}} \right)$, where α is a function of T .

Proof. Set $a(x) = 1/x$. To use Lemma 2, we check that $\int_0^1 a(x)^{-\frac{1}{k-1}} dx = \int_0^1 x^{\frac{1}{k-1}} dx < \infty$, and this is true for all $k \geq 1$. \square

4.1 The Optimal Function OPT_k

For each $k \geq 2$ we define the function $\text{OPT}_k \in \mathcal{F}((0, 1])$:

$$\text{OPT}_k(x, t) = \begin{cases} 1 & t < x/k \\ \left(\frac{x}{kt}\right)^{\frac{1}{k-1}} & x/k < t < 1/k \\ \frac{k}{k-1}(1-t)x^{\frac{1}{k-1}} & 1/k < t < 1 \\ 0 & t > 1 \end{cases} \quad (3)$$

See Figure 1 for an illustration of this function.

Claim 4. $\text{OPT}_k \in \mathcal{V}((0, 1])$, and every $N \in \mathcal{V}((0, 1])$ satisfies $\theta_N(k) \geq \theta_{\text{OPT}_k}(k)$. Also,

$$\theta_{\text{OPT}_k}(k) = \frac{3k-1}{k(k+1)}$$

The proof appears in Appendix A.3. It proceeds as suggested above: using Lemma 2 we know how each column of the optimal function looks like, we then use the column requirements to find them exactly. Eventually we calculate the double integral and get $\theta_{\text{OPT}_k}(k)$.

4.2 And now with ϵ

The following claim is what we need to conclude the upper bound proof and prove Claim 3:

Claim 5. *There is some real function f , s.t. $\lim_{\epsilon \rightarrow 0} f(\epsilon) = 1$, and every $N \in \mathcal{V}([\epsilon, 1])$ satisfies $\theta_N(k) \geq f(\epsilon) \cdot \theta_{OPT_k}(k)$.*

Proof. We construct a new function N' that will span the whole range of x 's from 0 to 1, with little change to $\theta_N(k)$. This will be done by slowing N down, and using what we saved in the column integrals to visit the x 's between 0 and ϵ using our optimal solution, running it fast enough so it does not incur a big difference in $\theta_N(k)$.

Fix some $a < 1$ to be determined later. Define:

$$N'(x, t) = \begin{cases} \text{OPT}_{k, \epsilon/(1-a)} & x \leq \epsilon \\ N_{1, 1/a} & x > \epsilon \end{cases}$$

Since the zoomed version of OPT_k here is defined on the x 's in $(0, \epsilon]$ and the zoomed N is on those in $[\epsilon, 1]$, we get that for all t :

$$C_{N'}(t) = C_{\text{OPT}_{k, \epsilon/(1-a)}}(t) + C_{N_{1, 1/a}}(t) = \epsilon C_{\text{OPT}_k} \left(\frac{1-a}{\epsilon} t \right) + C_N(at) \leq \epsilon \frac{1-a}{\epsilon} t + at = t$$

where we used Claim 2 and the fact that both OPT_k and N satisfy the column requirements.

N' can be made to be continuous with negligible cost to the integrals, as it is not continuous only on the seam between the two functions. Therefore, $N' \in \mathcal{V}((0, 1])$ and by Claim 4, we have $\theta_{N'}(k) \geq \theta_{\text{OPT}_k}(k)$. So, again by Claim 2:

$$\theta_{\text{OPT}_k}(k) \leq \theta_{N'}(k) = \epsilon \cdot \theta_{\text{OPT}_{k, \epsilon/(1-a)}}(k) + (1 - \epsilon) \cdot \theta_{N_{1, 1/a}}(k) = \frac{\epsilon}{1-a} \theta_{\text{OPT}_k}(k) + \frac{1-\epsilon}{a} \theta_N(k)$$

And therefore,

$$\frac{1-\epsilon}{a} \theta_N(k) \geq \left(1 - \frac{\epsilon}{1-a} \right) \theta_{\text{OPT}_k}(k) \implies \frac{\theta_N(k)}{\theta_{\text{OPT}_k}(k)} \geq \frac{a}{1-\epsilon} \left(1 - \frac{\epsilon}{1-a} \right)$$

Taking $a = 1 - \sqrt{\epsilon}$, the right size of the inequality becomes:

$$\frac{1-\sqrt{\epsilon}}{1-\epsilon} \left(1 - \frac{\epsilon}{\sqrt{\epsilon}} \right) = \frac{(1-\sqrt{\epsilon})^2}{1-\epsilon} = \frac{1-\sqrt{\epsilon}}{1+\sqrt{\epsilon}}$$

As ϵ goes to 0 this goes to 1. □

5 Matching Lower Bound

In this section we prove:

Theorem 3. *For all $k \geq 1$:*

$$\text{speedup}(k) \geq \frac{k(k+1)}{3k-1}$$

The case of $k = 1$ was dealt with in Section 2.3. For $k \geq 2$, denote by $A_{k,M}$, Algorithm 2 when set to run with parameter k and M . To prove Theorem 3, we prove:

Claim 6. *For every $k \geq 2$,*

$$\lim_{M \rightarrow \infty} \theta_{A_{k,M}}(k) \leq \frac{3k-1}{k(k+1)}$$

Proof. Denote the matrix of $A_{k,M}$ by N . A general observation is that:

$$N(x, t) = N(x, t-1) \cdot \Pr[x \text{ is not chosen at time } t]$$

Since the algorithm chooses uniformly from a set of unopened boxes at each stage, this is:

$$N(x, t) = N(x, t-1) \cdot \left(1 - \frac{1}{|\text{interval chosen from}| - (t-1)}\right)$$

Applying generously we get:

$$N(x, t) = \begin{cases} 1 & t < x/k \\ \prod_{i=\lceil x/k \rceil}^t \left(1 - \frac{1}{ik-i+1}\right) & x/k \leq t \leq M/k \\ \prod_{i=\lceil x/k \rceil}^{M/k} \left(1 - \frac{1}{ik-i+1}\right) \prod_{i=M/k+1}^t \left(1 - \frac{1}{M-i+1}\right) & M/k < t < M \\ 0 & t \geq M \end{cases}$$

where we assumed M is a multiple of k . Otherwise, take M large enough and round it up to be a multiple of k , the few extra boxes will not change the speed-up. Also, note that the value of $N(x, t)$ where x is not a multiple of k is the same as that of $N(x', t)$ where x' is the next multiple of k greater than x . We can therefore calculate the double sum of $\theta_N(k)$, when summing only over x 's that are multiples of k , and multiply the result by k . From here on we will assume x is a multiple of k .

We will need the following lemma:

Lemma 3. For integers $b \geq a \geq 1$, and $0 < \phi \leq 1$, $\prod_{i=a}^b \frac{i}{i+\phi} \leq \left(\frac{a}{b}\right)^\phi$.

Using properties of the Gamma function it is easy to see that the two sides of the equation are asymptotically equal, but this is not enough to prove our result as we need the inequality for small a and b as well. We prove this Lemma in Appendix A.4. Now:

$$\prod_{i=\frac{x}{k}}^t \left(1 - \frac{1}{ik-i+1}\right) = \prod_{i=\frac{x}{k}}^t \frac{ik-i}{ik-i+1} = \prod_{i=\frac{x}{k}}^t \frac{i}{i + \frac{1}{k-1}} \leq \left(\frac{x}{kt}\right)^{\frac{1}{k-1}}$$

And:

$$\begin{aligned} \prod_{i=\frac{M}{k}+1}^t \left(1 - \frac{1}{M-i+1}\right) &= \prod_{i=\frac{M}{k}+1}^t \frac{M-i}{M-i+1} = \frac{M - \frac{M}{k} - 1}{M - \frac{M}{k}} \cdots \frac{M-t}{M-t+1} \\ &= \frac{M-t}{M - \frac{M}{k}} = \frac{1 - \frac{t}{M}}{1 - \frac{1}{k}} = \frac{k}{k-1} \left(1 - \frac{t}{M}\right) \end{aligned}$$

So finally:

$$N(x, t) \leq \begin{cases} 1 & t < x/k \\ \left(\frac{x}{kt}\right)^{\frac{1}{k-1}} & x/k \leq t \leq M/k \\ \frac{k}{k-1} \left(1 - \frac{t}{M}\right) \left(\frac{x}{M}\right)^{\frac{1}{k-1}} & M/k < t < M \\ 0 & t \geq M \end{cases}$$

This is strangely familiar. Looking at (3) we notice that:

$$N(x, t) \leq \text{OPT}_k(x/M, t/M)$$

Calculating the sums of $\theta_N(k)$ is like approximating the double integral of θ_{OPT_k} , and as M goes to infinity this approximation can be as close as we wish. A fine detail is with the fact that we take only x 's that are multiples of k , but this creates a $1/k$ factor in the sum approximating the integral which is then cancelled by multiplying the total sum by k , as said above. \square

5.1 Algorithm with $O(\log(M))$ memory

As it is now, Algorithm 2 uses a lot of memory. Each searcher needs to remember all of the boxes it checked, so we need $O(M)$ memory in the worst case. However, if we are willing to lose a little in the speed-up, we can save a lot in memory. Algorithm 3 is exactly the same as Algorithm 2, except it doesn't remember

Algorithm 3 A memory efficient variant.

```

for  $i = 1$  to  $M/k$  do
    check a uniformly chosen box in  $\{1, \dots, i * k\}$ 
while there are unchecked boxes do
    check a uniformly chosen box out of all boxes.

```

where it's been. The only memory it uses is the counter, and so it needs only $O(\log(M))$ bits of memory. Denote by $B_{k,M}$, Algorithm 3 when set to run with parameter k and M . In Appendix A.5 we prove:

Claim 7. $\lim_{M \rightarrow \infty} \text{speedup}_{B_{k,M}}(k) \geq \frac{k}{3}$

As we can see, this algorithm, for large M and k will have essentially the same speed-up as that of the optimal algorithm while using exponentially less memory.

A Appendix

A.1 Proof of Claim 1

Claim 1. For any algorithm A that works on M boxes, there is a function $N \in \mathcal{V}([1, M+1])$ such that for all k , $\theta_N(k) \leq \theta_A(k)$.

Proof. Given A , take its matrix N_A , and define $N(x, t) = N_A(\lfloor x \rfloor, \lfloor t \rfloor)$. For any t :

$$\begin{aligned}
 C_N(t) &= \int_1^{M+1} 1 - N(x, t) dx = \int_1^{M+1} 1 - N_A(\lfloor x \rfloor, \lfloor t \rfloor) dx \\
 &= \sum_{x=1}^M 1 - N_A(x, \lfloor t \rfloor) = C_{N_A}(\lfloor t \rfloor) \leq \lfloor t \rfloor \leq t
 \end{aligned}$$

So N satisfies the column requirements. Next, let us upper-bound the inverse of the speed-up.

$$\begin{aligned}
 \theta_N(k) &= \frac{1}{M} \int_1^{M+1} \int_0^\infty \frac{1}{x} N_A(\lfloor x \rfloor, \lfloor t \rfloor)^k dt dx \leq \frac{1}{M} \int_1^{M+1} \int_0^\infty \frac{1}{\lfloor x \rfloor} N_A(\lfloor x \rfloor, \lfloor t \rfloor)^k dt dx \\
 &= \frac{1}{M} \sum_{x=1}^M \sum_{t=0}^\infty \frac{1}{x} N_A(x, t) = \theta_A(k)
 \end{aligned}$$

where the inequality is because $x \geq 1$ and $\lfloor x \rfloor \leq x$. N is a step function and not continuous but can be turned into a continuous function with as little change in integral as we wish - we can leverage what we gain in the inequality above to account for this change. \square

A.2 Proof of the Extended Lemma

Lemma 2. Fix $T \geq 0$. For continuous functions $a : (0, 1] \rightarrow \mathbb{R}^+$ and $f : (0, 1] \rightarrow [0, 1]$ where $\int_0^1 a(x)^{-1/(k-1)} dx$ is defined and $\int_0^1 1 - f(x) dx \leq T$, the minimum of $\int_0^1 a(x) f(x)^k dx$, ranging over all possible f 's is achieved when:

$$f(x) = \min \left(1, \alpha a(x)^{-\frac{1}{k-1}} \right)$$

Where α is a function of $a(\cdot)$ and T , and independent of x .

Proof. First, note that if $T \geq 1$ then the optimal solution is $f(x) = 0$ and then $\alpha = 0$. Otherwise, and this hold for Lemma 1 as well, we can assume that the constraint is actually an equality, as maximizing it will always decrease the target integral.

For each $0 < s < 1/2$, applying Lemma 1, there is some α_s such that among all functions $f : [s, 1] \rightarrow [0, 1]$, satisfying $\int_s^1 1 - f(x)dx \leq T - s$ (again, since $T < 1$, this is actually an equality), the function:

$$f_s(x) = \min(1, \alpha_s a(x)^{-1/(k-1)})$$

Minimizes $\int_s^1 a(x)f(x)^k dx$.

Extend each f_s to be also defined on $(0, s)$ by setting it to be 0 there. Since α_s is derived as the solution to:

$$\int_s^1 1 - \min(1, \alpha_s a(x)^{-\frac{1}{k-1}}) dx = T - s \implies \int_s^1 \min(1, \alpha_s a(x)^{-\frac{1}{k-1}}) dx = 1 - T$$

We can see that as s decreases, α_s decreases, and as the α_s are bounded below by 0, they converge to some α . Denote $f(x) = \min(1, \alpha a(x)^{-1/(k-1)})$.

Clearly, the f_s converge point-wise to f as s approaches 0. By the bounded convergence theorem we get:

$$\int_0^1 f(x)dx = \lim_{s \rightarrow 0} \int_0^1 f_s(x)dx = \lim_{s \rightarrow 0} \left(\int_0^s 0dx + \int_s^1 f_s(x)dx \right) = T$$

Since α_s decreases as s decreases, we get that for all $s < 1/2$:

$$a(x)f_s^k(x) = a(x) \min\left(1, \alpha_s a(x)^{-\frac{1}{k-1}}\right)^k \leq a(x) \left(\alpha_s a(x)^{-\frac{1}{k-1}}\right)^k \leq \alpha_{\frac{1}{2}}^k a(x)^{-\frac{1}{k-1}}$$

By our requirements on $a(\cdot)$ we know that $\int_0^1 a(x)^{-\frac{1}{k-1}} dx < \infty$, so by the dominated convergence theorem we can exchange integral and limit and get:

$$\int_0^1 a(x)f(x)^k dx = \lim_{s \rightarrow 0} \int_0^1 a(x)f_s(x)^k dx \tag{4}$$

Now, assume there is some g satisfying the conditions that is better than f . That is:

$$\int a(x)f(x)^k - \int a(x)g(x)^k > \delta > 0 \tag{5}$$

We take an s small enough so that (4) is more than $\delta/2$ close to the limit, and we get:

$$\int_0^1 a(x)f(x)^k - \int_0^1 a(x)g(x)^k < \frac{\delta}{2} + \int_0^1 a(x)f_s(x)^k - \int_0^1 a(x)g(x)^k$$

Yet we know that f_s is optimal on $[s, 1]$ and since it is 0 on $(0, s)$, g cannot be better. so this latter sum is at least $\delta/2$ contradicting (5). \square

A.3 Optimality of OPT_k

Claim 4. $\text{OPT}_k \in \mathcal{V}((0, 1])$, and any $N \in \mathcal{V}((0, 1])$ has $\theta_N(k) \geq \theta_{\text{OPT}_k}(k)$. Also,

$$\theta_{\text{OPT}_k}(k) = \frac{3k-1}{k(k+1)}$$

Proof. Corollary 2 gives for each fixed t the function $f_t(x)$ minimizing $\int_0^1 \frac{1}{x} f(x)^k dx$. Setting $\text{OPT}_k(x, t) = f_t(x)$ will give the minimal $\int_0^\infty \int_0^1 \frac{1}{x} \text{OPT}_k(x, t)^k dx dt$, which is the goal.

All we have to do is figure out $f_t(x)$. First, for $t > 1$, the optimal f_t is obviously $f_t(x) = 0$, since this satisfies the column requirement and contributes 0 to our target double integral. For $t < 1$, we have $\int_0^1 1 - f_t(x) dx \leq t$, and that $f_t(x) = \min(1, \alpha x^{1/(k-1)})$. So, given t , we figure out what its α is. For each t , denote by γ (a function of t) the smallest x where $N(x, t) = 1$, and in case this does not happen, set $\gamma = 1$:

$$\gamma = \min(1, \frac{1}{\alpha^{k-1}}) \quad (6)$$

for every $t < 1$, OPT_k satisfies the column requirement $C_{\text{OPT}_k}(t) \leq t$, and since it minimizes the target integral, it will actually be an equality:

$$1 - t = \int_0^1 \text{OPT}_k(x, t) dx$$

We have two cases:

1. for all t where $\gamma < 1$ this equation is:

$$1 - t = \int_0^\gamma \alpha x^{\frac{1}{k-1}} + \int_\gamma^1 1 dx \implies \gamma - t = \int_0^\gamma \alpha x^{\frac{1}{k-1}} dx$$

From (6), we get $\alpha = 1/\gamma^{1/(k-1)}$. Plugging this in:

$$\gamma - t = \int_0^\gamma \left(\frac{x}{\gamma}\right)^{\frac{1}{k-1}} dx = \gamma \int_0^1 x^{\frac{1}{k-1}} dx = \gamma \frac{1}{\frac{1}{k-1} + 1} = \frac{k-1}{k} \gamma$$

So:

$$\gamma - t = \gamma \frac{k-1}{k} \implies \gamma = kt$$

This means, that for all $t < \frac{1}{k}$, $\gamma < 1$ and then $\alpha = 1/(kt)^{1/(k-1)}$. For all other t , $\gamma = 1$.

2. for all t where $\gamma = 1$:

$$1 - t = \int_0^1 \alpha x^{\frac{1}{k-1}} dx = \alpha \frac{1}{\frac{1}{k-1} + 1} = \frac{k-1}{k} \alpha$$

and so:

$$\alpha = \frac{k}{k-1} (1 - t)$$

Putting all this together gives us $\text{OPT}_k(x, t)$ as in (3).

Now we can calculate $\theta_{\text{OPT}_k}(k)$:

$$\begin{aligned} \int_0^1 \int_0^1 \frac{1}{x} \text{OPT}_k(x, t)^k dx dt &= \int_0^{1/k} \left(\int_0^\gamma \frac{1}{x} (\alpha x^{\frac{1}{k-1}})^k dx \right) dt + \int_0^{1/k} \left(\int_\gamma^1 \frac{1}{x} 1^k dx \right) dt \\ &\quad + \int_{1/k}^1 \left(\int_0^1 \frac{1}{x} (\alpha x^{\frac{1}{k-1}})^k dx \right) dt \end{aligned}$$

We focus on each one of these:

1. Here we know that $\gamma = kt < 1$.

$$\int_0^\gamma \frac{1}{x} (\alpha x^{\frac{1}{k-1}})^k dx = \alpha^{k-1} \int_0^\gamma \alpha x^{\frac{1}{k-1}} dx = \alpha^{k-1}(\gamma - t) = \frac{1}{\gamma}(\gamma - t)$$

We get:

$$\frac{1}{kt}(kt - t) = \frac{k-1}{k}$$

The whole integral:

$$\int_0^{1/k} \frac{k-1}{k} = \frac{k-1}{k^2}$$

2. Here, still, $\gamma < 1$.

$$\int_\gamma^1 \frac{1}{x} dx = \log(1) - \log(\gamma) = -\log(\gamma)$$

Plugging our $\gamma = kt$ in, and calculating the whole integral:

$$\int_0^{1/k} -\log(kt) dt = -\frac{1}{k} \int_0^1 \log(t) dt = \frac{1}{k}$$

Last bit is because indefinite integral of $\log(x)$ is $x \log(x) - x$.

3. Here $\gamma = 1$.

$$\int_0^1 \frac{1}{x} (\alpha x^{\frac{1}{k-1}})^k dx = \alpha^{k-1} \int_0^1 \alpha x^{\frac{1}{k-1}} dx = \alpha^{k-1}(1 - t)$$

plugging in our $\alpha = \frac{k}{k-1}(1 - t)$ and calculating the whole integral:

$$\begin{aligned} \int_{1/k}^1 \left(\frac{k}{k-1}(1-t) \right)^{k-1} (1-t) dt &= \left(\frac{k}{k-1} \right)^{k-1} \int_0^{\frac{k-1}{k}} t^k dt \\ &= \left(\frac{k}{k-1} \right)^{k-1} \frac{1}{k+1} \left(\frac{k-1}{k} \right)^{k+1} = \frac{(k-1)^2}{k^2(k+1)} \end{aligned}$$

In total we get:

$$\begin{aligned} \frac{k-1}{k^2} + \frac{1}{k} + \frac{(k-1)^2}{k^2(k+1)} &= \frac{1}{k} \cdot \frac{(k-1)(k+1) + k(k+1) + (k-1)^2}{k(k+1)} \\ &= \frac{1}{k} \cdot \frac{k^2 - 1 + k^2 + k + k^2 - 2k + 1}{k(k+1)} = \frac{1}{k} \cdot \frac{3k^2 - k}{k(k+1)} = \frac{3k-1}{k(k+1)} \end{aligned}$$

□

A.4 Proof of the Gamma Function Property

Lemma 3. For integers $b \geq a \geq 1$, and $0 < \phi \leq 1$,

$$\prod_{i=a}^b \frac{i}{i+\phi} \leq \left(\frac{a}{b} \right)^\phi$$

Proof. By induction on a (somehow on b it doesn't work..). If $b = a$, then we should show $a/(a + \phi) \leq 1$, which is true. We therefore assume that the result holds for $a + 1$ and prove it for a :

$$\prod_{i=a}^b \frac{i}{i + \phi} = \frac{a}{a + \phi} \cdot \prod_{i=a+1}^b \frac{i}{i + \phi} \leq \frac{a}{a + \phi} \cdot \left(\frac{a+1}{b}\right)^\phi$$

We want to show:

$$\begin{aligned} \frac{a}{a + \phi} \cdot \left(\frac{a+1}{b}\right)^\phi &\leq \left(\frac{a}{b}\right)^\phi \iff \frac{a}{a + \phi} \leq \left(\frac{a}{a+1}\right)^\phi \\ &\iff \left(\frac{a + \phi}{a}\right)^{\frac{1}{\phi}} \geq \left(\frac{a+1}{a}\right) \end{aligned}$$

Take $b = \frac{1}{a} < 1$ and $x = \frac{1}{\phi} \geq 1$ the above is equivalent to:

$$\left(1 + \frac{b}{x}\right)^x \geq (1 + b)$$

If we show that the left side is increasing with x when $x \geq 1$ then we are done. We take the derivative (using an internet site):

$$\left(1 + \frac{b}{x}\right)^x \cdot \left(\ln\left(1 + \frac{b}{x}\right) - \frac{b}{x(1 + \frac{b}{x})}\right)$$

This is positive if

$$\left(1 + \frac{b}{x}\right) \ln\left(1 + \frac{b}{x}\right) > \frac{b}{x}$$

Take $y = \frac{b}{x} \leq 1$. We want to show that

$$(1 + y) \ln(1 + y) > y$$

We use the equality

$$\ln(1 + y) = \int_0^y \frac{1}{1 + t} dt$$

So

$$(1 + y) \ln(1 + y) = \int_0^y \frac{1 + y}{1 + t} dt > \int_0^y 1 dt = y$$

as desired. □

A.5 Speedup of Algorithm 3

Claim 7.

$$\lim_{M \rightarrow \infty} \text{speedup}_{B_{k,M}}(k) \geq \frac{k}{3}$$

Proof. By the same reasoning as the original algorithm, the matrix of this one is:

$$N(x, t) = \begin{cases} 1 & t < x/k \\ \prod_{i=\lceil x/k \rceil}^t \left(1 - \frac{1}{ik}\right) & x/k \leq t \leq M/k \\ \prod_{i=\lceil x/k \rceil}^{M/k} \left(1 - \frac{1}{ik}\right) \cdot \left(1 - \frac{1}{M}\right)^{t-M/k} & t > M/k \end{cases}$$

Using Lemma 3:

$$\prod_{i=x/k}^t \left(1 - \frac{1}{ik}\right) \leq \prod_{i=x/k}^t \left(1 - \frac{1}{ik+1}\right) = \prod_{i=x/k}^t \frac{ik}{ik+1} = \prod_{i=x/k}^t \frac{i}{i + \frac{1}{k}} \leq \left(\frac{x}{tk}\right)^{\frac{1}{k}}$$

Using it and ignoring all the rounding, as in the analysis of Algorithm 2:

$$N(x, t) \leq \begin{cases} 1 & t < x/k \\ \left(\frac{x}{tk}\right)^{\frac{1}{k}} & x/k \leq t \leq M/k \\ \left(\frac{x}{M}\right)^{\frac{1}{k}} \cdot \left(1 - \frac{1}{M}\right)^{t-M/k} & t > M/k \end{cases}$$

To upper bound $\theta_N(k) = \frac{1}{M} \sum_{x=1}^M \frac{1}{x} \sum_{t=0}^{\infty} N(x, t)^k$, we calculate each area of the sum separately. In what follows, we use the fact that M is large for a few approximations.

1. The 1's contribute:

$$\frac{1}{M} \sum_{x=1}^M \frac{1}{x} \sum_{t=0}^{x/k} 1^k = \frac{1}{M} \sum_{x=1}^M \frac{1}{k} = \frac{1}{k}$$

2. The second part contributes:

$$\frac{1}{M} \sum_{x=1}^M \frac{1}{x} \sum_{t=x/k}^{M/k} \frac{x}{tk} = \frac{1}{M} \sum_{x=1}^M \sum_{t=x/k}^{M/k} \frac{1}{tk} \approx \frac{1}{Mk} \sum_{x=1}^M \sum_{t=x}^M \frac{1}{t} = \frac{1}{Mk} \sum_{x=1}^M \sum_{t=x}^M \frac{1}{t} = \frac{1}{Mk} M = \frac{1}{k}$$

where the double sum is equal to M , by noting that 1 appears there once, $\frac{1}{2}$ appear twice, and so on.

3. The last part is:

$$\begin{aligned} \frac{1}{M} \sum_{x=1}^M \frac{1}{x} \left(\frac{x}{M}\right) \sum_{t=M/k}^{\infty} \left(1 - \frac{1}{M}\right)^{\left(t - \frac{M}{k}\right)k} &= \frac{1}{M^2} \sum_{x=1}^M \sum_{t=0}^{\infty} \left(1 - \frac{1}{M}\right)^{tk} \\ &= \frac{1}{M} \frac{1}{1 - \left(1 - \frac{1}{M}\right)^k} \approx \frac{1}{M} \frac{1}{1 - \left(1 - \frac{k}{M}\right)} = \frac{1}{k} \end{aligned}$$

Putting it all together, we get the result. □

References

- [1] Noga Alon, Chen Avin, Michal Koucky, Gady Kozma, Zvi Lotker, and Mark R. Tuttle. Many Random Walks Are Faster Than One. In *Proceedings of the Twentieth Annual Symposium on Parallelism in Algorithms and Architectures*, SPAA '08, pages 119–128, New York, NY, USA, 2008. ACM.
- [2] R.A. Baezayates, J.C. Culberson, and G.J.E. Rawlins. Searching in the Plane. *Inf. Comput.*, 106(2):234–252, October 1993.
- [3] Anatole Beck. On the linear search problem. *Israel Journal of Mathematics*, 2(4):221–228, 1963.
- [4] Richard Bellman. An Optimal Search. *SIAM Review*, 5(3):274–274, 1963.
- [5] Colin Cooper, Alan M. Frieze, and Tomasz Radzik. Multiple Random Walks in Random Regular Graphs. *SIAM J. Discrete Math.*, 23(4):1738–1761, 2009.

- [6] Erik D. Demaine, Sándor P. Fekete, and Shmuel Gal. Online searching with turn cost. *Theor. Comput. Sci.*, 361(2-3):342–355, 2006.
- [7] Klim Efremenko and Omer Reingold. How Well Do Random Walks Parallelize? In Irit Dinur, Klaus Jansen, Joseph Naor, and José Rolim, editors, *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques: 12th International Workshop, APPROX 2009, and 13th International Workshop, RANDOM 2009, Berkeley, CA, USA, August 21-23, 2009. Proceedings*, pages 476–489. Springer Berlin Heidelberg, Berlin, Heidelberg, 2009.
- [8] Robert Elsässer and Thomas Sauerwald. Tight bounds for the cover time of multiple random walks. *Theor. Comput. Sci.*, 412(24):2623–2641, 2011.
- [9] Yuval Emek, Tobias Langner, David Stolz, Jara Uitto, and Roger Wattenhofer. How many ants does it take to find the food? *Theor. Comput. Sci.*, 608:255–267, 2015.
- [10] Yuval Emek, Tobias Langner, Jara Uitto, and Roger Wattenhofer. Solving the ANTS Problem with Asynchronous Finite State Machines. In *Automata, Languages, and Programming - 41st International Colloquium, ICALP 2014, Copenhagen, Denmark, July 8-11, 2014, Proceedings, Part II*, pages 471–482, 2014.
- [11] Ofer Feinerman and Amos Korman. Memory Lower Bounds for Randomized Collaborative Search and Implications for Biology. In *Distributed Computing - 26th International Symposium, DISC 2012, Salvador, Brazil, October 16-18, 2012. Proceedings*, pages 61–75, 2012.
- [12] Ofer Feinerman, Amos Korman, Zvi Lotker, and Jean-Sébastien Sereni. Collaborative search on the plane without communication. In *ACM Symposium on Principles of Distributed Computing, PODC '12, Funchal, Madeira, Portugal, July 16-18, 2012*, pages 77–86, 2012.
- [13] Pierre Fraigniaud, Amos Korman, and Yoav Rodeh. Parallel Exhaustive Search without Coordination. *CoRR*, abs/1511.00486, 2015. To appear in STOC 2016.
- [14] Shmuel Gal. Minimax Solutions for Linear Search Problems. *SIAM Journal on Applied Mathematics*, 27(1):17–30, 1974.
- [15] Ming-Yang Kao, John H. Reif, and Stephen R. Tate. Searching in an Unknown Environment: An Optimal Randomized Algorithm for the Cow-path Problem. In *Proceedings of the Fourth Annual ACM-SIAM Symposium on Discrete Algorithms, SODA '93*, pages 441–447, Philadelphia, PA, USA, 1993. Society for Industrial and Applied Mathematics.
- [16] Richard M. Karp, Michael E. Saks, and Avi Wigderson. On a Search Problem Related to Branch-and-Bound Procedures. In *27th Annual Symposium on Foundations of Computer Science, Toronto, Canada, 27-29 October 1986*, pages 19–28, 1986.
- [17] Tobias Langner, Jara Uitto, David Stolz, and Roger Wattenhofer. Fault-Tolerant ANTS. In *Distributed Computing - 28th International Symposium, DISC 2014, Austin, TX, USA, October 12-15, 2014. Proceedings*, pages 31–45, 2014.
- [18] Christoph Lenzen, Nancy A. Lynch, Calvin C. Newport, and Tsvetomira Radeva. Trade-offs between selection complexity and performance when searching the plane without communication. In *ACM Symposium on Principles of Distributed Computing, PODC '14, Paris, France, July 15-18, 2014*, pages 252–261, 2014.